Journal of Forestry Research (2009) 20(4): 285–292
DOI 10.1007/s11676-009-0050-y

RESEARCH PAPER

Small-area estimation of forest stand structure in Jalisco, Mexico

Robin M. Reich[1], Celedonio Aguirre-Bravo[2]

[1]*Department of Forest, Rangeland and Watershed Stewardship, Colorado State University, Fort Collins, Colorado 80523 USA*
[2]*Rocky Mountain Research Station, USDA Forest Service, Fort Collins, Colorado 80523 USA*

Abstract: Natural resource statistics are often unavailable for small ecological or economic regions and policymakers have to rely on state-level datasets to evaluate the status of their resources (i.e., forests, rangelands, grasslands, agriculture, etc.) at the regional or local level. These resources can be evaluated using small-area estimation techniques. However, it is unknown which small area technique produces the most valid and precise results. The reliability and accuracy of two methods, synthetic and regression estimators, used in small-area analyses, were examined in this study. The two small-area analysis methods were applied to data from Jalisco's state-wide natural resource inventory to examine how well each technique predicted selected characteristics of forest stand structure. The regression method produced the most valid and precise estimates of forest stand characteristics at multiple geographical scales. Therefore, state and local resource managers should utilize the regression method unless appropriate auxiliary information is not available.

Key words: forest structure; regression estimator; synthetic estimator; spatial model; stratified random sampling; satellite imagery; inventory and monitoring

Introduction

Health of ecosystems is a matter of national security for most industrial economies. Yet conventional approaches to land resource inventory and monitoring are being significantly challenged by new scientific needs and social environmental concerns. One solid example of this is the State of Jalisco's ecosystem resource inventory and monitoring program, known in Mexico under the acronym of IMRENAT (Flores-Garnica et al. 2007). From the beginning, policy makers and stake holders in Jalisco called for an inventory and monitoring program that was responsive to the spatiotemporal dependencies of natural resources as influenced by natural and human processes. In addition, the design not only focused on classical statistical population parameters (Schreuder et al. 1993) for commercially valuable resources, but also had to account for estimates of other important ecological indicators that are fundamental to support decision making processes for ecosystem sustainability at multiple geographical and organizational scales. Consequently, IMRENAT was designed to meet these needs and concerns. IMRENAT was implemented in the year 2006; and results presented in early summer of 2007 (SEDER-FIPRODEFO 2007). Design specifications and assessment of its statistical performance have been reported (Reich et al. 2008b) for key land resource variables (e.g. canopy closure, basal area, cubic volume of commercial trees, and forest tree biomass) across bioclimatic regions.

Federal and state funds in Mexico have been used to create many new environmental and land resource management initiatives. Stakeholders want to know the current status of the natural resources in their region, county, watershed, and land units. However, current sources of data do not provide information at these organizational scales. At the state level IMRENAT is the primary source of information as it provides annual estimates of the status of natural resources (i.e., forests, rangelands, grasslands, agriculture, etc.). Estimates of these and other land resource variables are needed for several levels of geopolitical and ecological administrative units, what in Mexico is termed *Territory Land Resource Management* (Ordenamiento Territorial). The demand for information at the local level has greatly increased, and producing local area statistics has emerged as one of the state's most difficult and pressing statistical problem. SEDER-FIPRODEFO is responsible for providing statistics on the status of the natural resources for use in ecological sustainability research, resource management and planning and assessments of land resources to meet regional and local economic needs.

IMRENAT was designed to provide reliable estimates at the state level. Even within the broad climatic regions within the state, sample sizes are large enough to produce direct estimates

Received: 2009-04-20; Accepted: 2009-05-22

The online version is available at http://www.springerlink.com
Biography: Dr. Robin M. Reich (1954-), male, is a professor of forest biometrics/spatial statistics in the Department of Forest, Rangeland and Watershed Stewardship at Colorado State University. Dr. Reich is an expert in the application of spatial statistics in designing natural resource inventories and ecosystem modeling. Tel: 970-491-6980; Fax: 970-491-6754; Email: robin@warnercnr.colostate.edu
Responsible editor: Chai Ruihai

with a reasonable level of precision. However, for many small geographical regions within the state, sample sizes are often too small to produce reliable estimates. When a small area contains sample data, this permits the use of models with specific small area effects, allowing for more accurate estimates of the parameters of interest at the small area level (Fay and Herriot 1979; Gosh and Rao 1994; Rao 1999; Pfeffermann 1999; Lehtonen and Pahkinen 2004). If no sample data is available, indirect estimators, such as synthetic estimators, are required to produce reliable estimates (Rao 2003; Ghosh and Rao 1994). The technique of synthetic estimation involves applying national or regional estimates of the characteristic being measured for specific population subgroups to the small area's population composition (Laake 1978). The simplest form of synthetic estimation requires computation of a weighted average of the mean values of the characteristics in the subgroups with weights that are proportional to the distribution of the subgroups in the small-area population. A more general approach involves regression analysis. In this approach, national or regional data are used to estimate a regression equation which relates the independent variables, which define the population subgroups, to the characteristic of interest. The values of the regression variables for the small area are then used in the equation to obtain estimates of the characteristic for that small area.

This paper briefly describes two techniques which can be used to produce small area estimates of forest structure. An overview of IMRENAT, which focuses on aspects of the survey design that, in turn, affect small area estimation will be presented. Each of the techniques is then used with data from IMRENAT to produce estimates of selected forest stand characteristics for several geopolitical regions in the state. Strengths and weaknesses of the techniques are discussed in the context of choosing the best approach for the small area system.

Methods

Notation

Consider a finite population $\Omega = (1,\ldots,N)$ consisting of N identifiable units. A stratification variable $\widetilde{J}$ is used to subdivide the population into H strata with N_h units in stratum h from which a simple random sample of n_h units are selected without replacement. The total sample size $n = \sum_{h=1}^{H} n_h$ and population size $N = \sum_{h=1}^{H} N_h$. Associated with the ith unit of the hth stratum are two variables y_{hi}, the variable of interest and x_{hi}, a covariate that is available for all units in the population. For the hth stratum let $W_h = N_h / N$ be the stratum weight, $f_h = n_h / N_h$ the sampling fraction, $\bar{y}_h$, $\bar{x}_h$, $\bar{Y}_h$, $\bar{X}_h$ the y- and x- sample and population means respectively. Assume $\bar{X} = \sum_{h=1}^{H} W_h \bar{X}_h$ is known. Let $\Omega_d = (1,\ldots,D)$ represent a small area population consisting of N_d units $\left(\bigcup_{d=1}^{D} \Omega_d = \Omega; \sum_{d=1}^{D} N_d = N\right)$ which are members of the $K \le H$ strata each consisting of N_{dh} units $\left(\sum_{h=1}^{K\le H} N_{dh} = N_d; \sum_{d=1}^{D} N_{dh} = N_h\right)$ with stratum weights $W_{dh} = N_{dh} / N_d$ and population mean $\bar{X}_{dh}$.

Synthetic estimator

The synthetic estimate for a small area Ω_d is the sum of the weighted average of the stratum means, $\bar{y}_h$ across all strata within the small area, where the weight is the proportion of the population of the small area Ω_d that is in each stratum. That is,

$$\hat{\bar{y}}_d = \sum_{h=1}^{K\le H} W_{dh} \bar{y}_h \tag{1}$$

with estimated variance

$$\hat{V}\left(\hat{\bar{y}}_d\right) = \sum_{h=1}^{K\le H} W_{dh}^2 \left(\frac{s_h^2}{n_h}\right) \tag{2}$$

where s_h^2 is the sample variance for stratum h. The appropriateness of this approach inherently depends upon the extent to which the criterion used to stratify the population is related to the parameter of interest.

Regression estimator

The regression estimator is based on a regression equation using a predictor variable as the independent variable and sample data for the variable of interest as the dependent variable. Estimates of the dependent variable are computed at the stratum level from the field data. Using the population means of the independent variable for each stratum a regression equation to predict $\bar{y}_h$ is:

$$\hat{\bar{y}}_h = a + b\bar{X}_h . \tag{3}$$

To adjust the relative importance of any one stratum (h) to the overall mean and to reduce any bias in survey estimates due to local differences in the response variable, the stratum means were multiplied by the stratum weights ($z_h = W_h \bar{y}_h$, $v_h = W_h \bar{X}_h$) prior to fitting the regression equation (DuMouchel and Duncan 1983). The weighted stratum means were used to define the regression estimator (Cochran 1977):

$$\hat{z}_{dh} = \bar{z} + b\left(v_{dh} - \bar{v}\right). \tag{4}$$

To obtain an estimate for the small area Ω_d , estimates from the regression equation are obtained for each stratum and summed

$$\sum_{h=1}^{H} z_{dh} = H\bar{z} + b\left(\sum_{h=1}^{H} v_{dh} - H\bar{v}\right). \qquad (5)$$

Letting $H\bar{z} = \sum_{h=1}^{H} W_h \bar{y}_h = \hat{\bar{y}}_{st}$ and $H\bar{v} = \sum_{h=1}^{H} W_h \bar{X}_h = \bar{X}$ the small area estimator $\sum_{h=1}^{H} z_{dh} = \sum_{h=1}^{H} W_{dh}\hat{\bar{y}}_h = \hat{\bar{y}}_d$ is obtained by substituting $\bar{X}_d$ for $\sum_{h=1}^{H} v_{dh} = \sum_{h=1}^{H} W_{dh}\bar{X}_{dh}$, giving

$$\hat{\bar{y}}_d = \hat{\bar{y}}_{st} + b\left(\bar{X}_d - \bar{X}\right). \qquad (6)$$

Thus, the estimate of the population parameter for the small area Ω_d is equal to the estimated population mean ($\hat{\bar{y}}_{st}$) from the state-wide inventory adjusted for differences between the population means of the independent variable at the small area and state level.

The prediction variance for stratum h in the small-area of interest, Ω_d is given by (Cochran 1977)

$$\hat{V}(\hat{z}_{dh}) = MSE\left(1 + \frac{1}{H} + \frac{(v_{dh} - \bar{v})^2}{\sum_{h=1}^{L}(v_h - \bar{v})^2}\right) \qquad (7)$$

where

$$MSE = \frac{1}{(n-2)}\left(\sum_{h=1}^{H}(z_h - \bar{z})^2 - b^2\sum_{h=1}^{H}(v_h - \bar{v})^2\right) \qquad (8)$$

Summing across all K ≤ L strata the prediction variance for $\hat{\bar{y}}_d$ becomes

$$\hat{V}\left(\hat{\bar{y}}_d\right) = MSE\left(K + \frac{K}{H} + \frac{\sum_{h=1}^{K<H}(v_{dh} - \bar{v})^2}{\sum_{h=1}^{H}(v_h - \bar{v})^2}\right) \qquad (9)$$

where $\bar{v} = \frac{1}{H}\sum_{h=1}^{H} W_h \bar{X}_h$, $v_{dh} = W_{dh}\bar{X}_{dh}$ and v_h is as defined previously.

The motivation for use of this type of estimator is that if the set of independent variables are easily obtainable for the small area and if the relationship between the dependent and independent variables is strong, then estimates of good quality might be produced at relatively low cost.

Example

We compare the two estimators (i.e., synthetic and regression) for forest stand characteristics and their estimated variances for 12 economic regions in the State of Jalisco; the eight counties in the Sierra Ocidente economic region in western Jalisco as well as for two small watersheds and a land tenure unit within this economic region (Fig. 1).

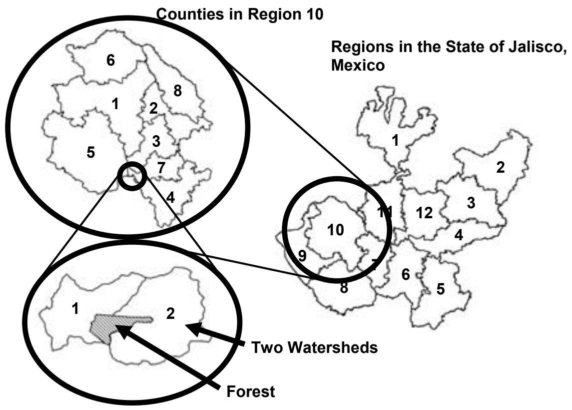

Fig. 1 Location of twelve economic regions in the State of Jalisco, Mexico; eight counties in the Sierra Ocidente economic region in western Jalisco and two small watersheds and a land tenure unit within this economic region.

A two-way nested design was used to stratify the state for the purpose of the inventory and monitoring program (IMRENAT). The state was first stratified by climatic zones (Reich et al. 2008a) which corresponded to broad vegetation types (i.e., semi-arid, temperate and tropical) in the state. Each of the three climatic zones was further stratified into forested and non-forested areas resulting in a total of six strata. Nested within each of the six strata, satellite imagery (Landsat-7 ETM+) was used to identify ten spectral classes to account for the variability in the vegetation cover, resulting in a total of H = 60 strata. A total of n = 1442, 30 m × 30 m primary sampling units were randomly allocated to strata based on the economic importance of the region and whether it was forested or non-forested (Table 1).

Variables selected for analysis included percent canopy closure, basal area ($m^2 \cdot ha^{-1}$), cubic volume ($m^3 \cdot ha^{-1}$) and forest tree biomass ($tones \cdot ha^{-1}$). Trees (≥12.5 cm DBH) were measured for diameter at breast height (DBH) and total tree height (m) and recorded by species on each sample plot. Tree diameters were used to estimate basal area on each sample plot. A spherical densiometer was used to estimate the average percent canopy closure on each sample plot. Tree diameters and total tree height were used to estimate cubic volumes and above ground biomass using regression equations provided by the State. Biomass was estimated for all trees, while volumes were only estimated for commercial species. Volumes for non-commercial species were set to zero.

Table 1. Distribution of sample plots in the network of permanent plots in Jalisco, Mexico by bioclimatic zone, the forest/non-forest stratum, and spectral class.

	Semi-Arid		Temperate		Tropical	
Spectral Class	Non-forested	Forested	Non-forested	Forested	Non-forested	Forested
1	9	16	0	27	1	22
2	5	50	12	44	15	34
3	8	28	32	40	7	25
4	4	35	24	58	8	31
5	9	24	29	78	16	28
6	6	26	20	40	11	41
7	3	19	38	33	10	28
8	11	30	11	66	14	34
9	8	20	24	64	14	53
10	2	16	8	42	7	24
Total	65	264	198	492	103	320

The small area regression estimator requires the availability of an independent variable that is correlated to the variable of interest and is available for all sample units in the population to allow the computation of population means for individual stratum for the population and small area. This information was provided in the form of a set of spatial models developed as part of IMRENAT (Reich et al. 2008b). The models were used to predict the spatial distribution of previously selected variables throughout the state at the same spatial resolution as the field data and satellite imagery used to stratify the state (Fig. 2). Details of the modeling process can be found in Reich et al. (2008b). The modeling process involved using multiple linear regressions to describe the large-scale spatial variability, while a tree-based stratified design was used to describe the small-scale variability associated with site-specific variability in forest stand structure. Independent variables used in the models included various Landsat-7 ETM+ bands, climatic and topographic data. The fitted models described 63% of the observed variability in canopy closure, 62% for basal area, 79% for volume and 82% for biomass used in their study.

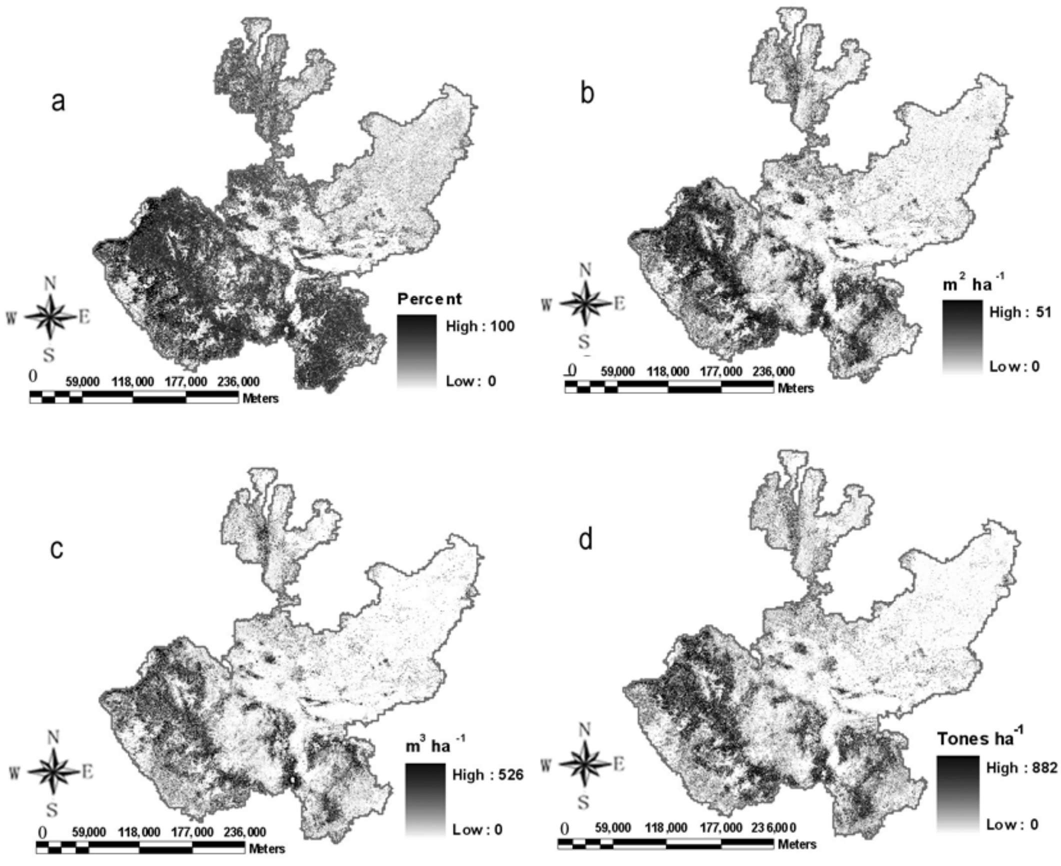

Fig. 2 Spatial distribution of a) canopy closure (%), b) tree basal area ($m^2 \cdot ha^{-1}$), c) cubic volumes ($m^3 \cdot ha^{-1}$) of commercially important tree species and d) above ground tree biomass ($tonnes \cdot ha^{-1}$) in the State of Jalisco, Mexico.

Within each of the H=60 strata, estimates of the mean and sample variance were calculated for each variable from the field data. Stratum means from the spatial estimates ($\overline{X}_h$) were obtained by averaging these estimates over all N_h sample units within a stratum. Stratum sizes were defined as the number of sample units associated with each stratum. This resulted in 60 paired sets of stratum means from the 60 strata for each dependent variable (i.e., field data) and independent variable (i.e., spa-

tial models). The stratum means were then weighted by their stratum weights. Ordinary least squares was then used to estimate the parameters of the individual regression equations relating the weighted stratum means from the spatial models to the weighted stratum means from the state-wide inventory. For each of the small areas identified in Fig. 1, estimates of the mean and standard errors were obtained using the synthetic and regression estimators. Estimates were compared with respect to their precision and a two-sample t-test was used to test for significant differences in estimated means. Small area estimates, when aggregated over all small areas, were further evaluated on their consistency with state level estimates.

Results

Figure 3 shows the standardized stratum means for canopy closure, basal area, volume and biomass for the 60 strata in the state, along with the stratum weights. Stratum means were standardized by subtracting the median value from the individual stratum means and dividing by the range in stratum means. The figure shows the relationship between the stratum means for the forest characteristics and the spectral classes used to stratify the state. Since all four variables are correlated they exhibited similar trends within strata. Direct estimates of the forest stand characteristics for the three climatic regions are provided in Table 2 along with their estimated variances and percent sampling errors.

Table 2. Direct estimates of forest stand characteristics for the three bioclimatic regions in the State of Jalisco, Mexico.

Stand Characteristic	Sample Size	Min.	Mean	Max.	Variance of Mean	%SE
Semi-Arid (Area = 1,614,167 ha)						
Canopy Closure (%)	329	0	21.3	97	0.89	8.8
Basal Area ($m^2 \cdot ha^{-1}$)		0	3.6	107	0.18	23.8
Volume ($m^3 \cdot ha^{-1}$)		0	19.9	877	11.5	34.0
Biomass ($tonnes \cdot ha^{-1}$)		0	29.7	1776	41.0	43.2
Temperate (Area = 4,738,465 ha)						
Canopy Closure (%)	690	0	34.8	100	0.85	5.3
Basal Area ($m^2 \cdot ha^{-1}$)		0	5.0	59	0.05	8.7
Volume ($m^3 \cdot ha^{-1}$)		0	29.8	658	3.1	11.8
Biomass ($tonnes \cdot ha^{-1}$)		0	36.4	483	3.0	9.6
Tropical (Area = 1,493,114 ha)						
Canopy Closure (%)	423	0	45.5	100	2.1	6.4
Basal Area ($m^2 \cdot ha^{-1}$)		0	9.0	142	0.33	12.8
Volume ($m^3 \cdot ha^{-1}$)		0	50.5	1275	18.2	16.9
Biomass ($tonnes \cdot ha^{-1}$)		0	63.3	749	17.2	13.1

The relationship between the weighted stratum means obtained from the spatial models and the state-wide inventory for the selected variables are displayed in Fig. 4 with fit statistics in Table 3. Except for the volume equation the estimated slopes did not differ significantly from 1 at the 0.05 level of significance. Also, the estimated intercepts did not differ significantly from 0 at the 0.05 level of significance. These results suggest a close agreement between the weighted stratum means from the spatial models and the field data. Histograms of the residuals and scatter plots of the residuals versus the fitted values showed no gross deviations from the assumptions of the linear model.

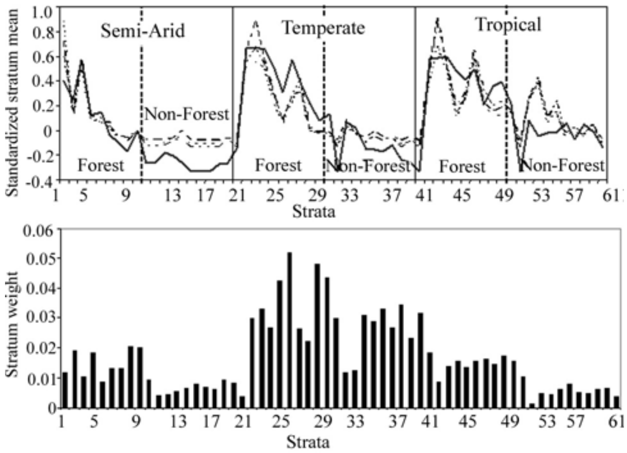

Fig. 3 Standardized stratum means for canopy closure (%), basal area ($m^2 \cdot ha^{-1}$), cubic volumes ($m^3 \cdot ha^{-1}$) of commercially important tree species, and above ground tree biomass ($tonnes \cdot ha^{-1}$) for the 60 strata in the state, along with the stratum weights. Stratum means were standardized by subtracting the median value from the individual stratum means and dividing by the range in stratum means.

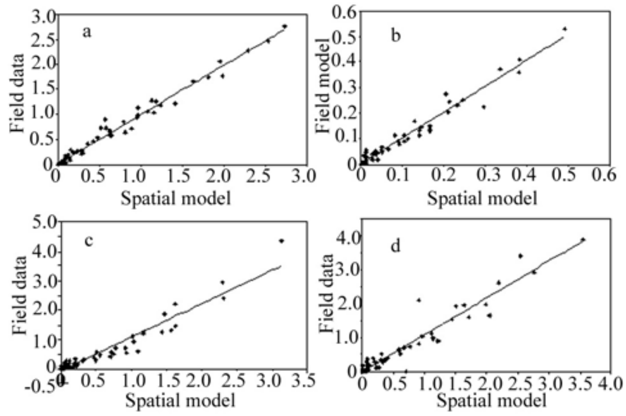

Fig. 4 Scatter plots of the relationships between the weighted stratum means from the state-wide inventory and the weighted stratum means from the spatial models for (a) canopy closure (%), (b) tree basal area ($m^2 \cdot ha^{-1}$), (c) cubic volumes ($m^2 \cdot ha^{-1}$) of commercially important species and (d) above ground tree biomass (tonnes/ha) in the state of Jalisco, Mexico. The solid lines are the regression lines for the individual variables.

A comparison of the two estimators in estimating basal area ($m^2 \cdot ha^{-1}$) in the 12 economic regions is summarized in Table 4. The table contains estimates of the mean, variances and percent sampling errors for each economic region. Also included in the tables is the ratio of the variance for the synthetic estimator to the variance of the regression estimator. This ratio will be greater than one when the regression estimator is more precise than the synthetic estimator. A t-statistic for testing the null hypothesis of no difference in the estimated means is also provided.

Table 3. Estimated regression coefficients relating the weighted stratum means obtained from the spatial models to the weighted stratum means obtained from the state-wide inventory for the L=60 strata.

Model	n	$\hat{\beta}_0$	$se(\hat{\beta}_0)$	$\hat{\beta}_1$	$se(\hat{\beta}_1)$	R^2	p-value[1] $H_0 : \beta_1 = 1$
Canopy Closure (%)	60	0.014	0.016	0.988	0.017	0.982	0.490
Basal Area (m^2/ha)	60	0.002	0.003	1.018	0.025	0.965	0.469
Volume (m^3/ha)	60	-0.066	0.035	1.144	0.035	0.928	< 0.001
Biomass (tones/ha)	60	0.006	0.039	1.078	0.040	0.926	0.053

[1] p-value testing the null hypothesis that the slope of the regression equation is equal to one

Table 4. Comparisons of the synthetic and regression estimators for estimating basal area ($m^2·ha^{-1}$) at the regional level in the State of Jalisco, Mexico.

		Synthetic Estimator			Regression Estimator				
Area	Name	Mean	Variance	%SE	Mean	Variance	%SE	Rel.Eff.¶	t-Statistic§
1	Norte	4.8	0.11	14.1	4.2	0.03	8.3	3.8	1.59
2	Altos Norte	1.9	0.02	15.7	1.3	0.02	22.5	1.2	3.27*
3	Altos Sur	2.3	0.02	11.6	1.5	0.02	18.6	0.8	3.77*
4	Cienega	2.2	0.01	11.4	1.5	0.02	15.2	0.8	3.53*
5	Sureste	7.8	0.08	7.5	7.6	0.03	4.7	2.6	0.38
6	Sur	6.8	0.10	9.4	6.8	0.03	5.3	3.2	0.16
7	Sierra de Amula	4.6	0.04	8.2	4.8	0.03	7.3	1.2	-0.56
8	Costa Sur	9.2	0.20	9.7	9.5	0.03	3.7	6.2	-0.80
9	Costa Norte	9.1	0.34	12.8	8.7	0.03	4.0	11.3	0.51
10	Sierra Ocidente	8.8	0.13	8.2	10.7	0.03	3.4	3.9	-4.84*
11	Valles	4.2	0.04	8.9	4.1	0.03	8.4	1.2	0.49
12	Centro	3.0	0.02	10.0	2.7	0.03	12.7	0.8	1.27

¶ Ratio of the variance of the synthetic estimator to the variance of the regression estimator.

§ Two-sample t-test for differences in the estimated population means. * = significant at the 0.05 level.

The synthetic estimator had a tendency to overestimate in the semi-arid region (regions: 2, 3, 4, 12) and underestimate in the tropical region (regions: 8, 9) relative to the regression estimator. This trend was consistent with the expectations of managers familiar with the forest resources in these economic regions. Significant differences between the two sets of estimates occurred primarily in the semi-arid region. More significant differences were observed between estimates of canopy closure and volume than estimates for basal area and biomass. The regression estimator was more precise in estimating canopy closure across all economic regions. In the semi-arid region the synthetic estimator was more precise in estimating basal area, volume and biomass. The synthetic estimator was also more precise than the direct estimator in the semi-arid region (Table 2). Both the direct and synthetic estimators had similar levels of precision in the temperate and tropical regions.

To further evaluate the regression estimator it was used to estimate the forest characteristics for the eight counties in the Sierra Ocidente economic region in western Jalisco and two small watersheds and a land tenure unit within this economic region. The results for cubic volumes ($m^3·ha^{-1}$) are summarized in Table 5. Included in the table are the size of each region, estimates of the mean, variances, lower and upper 0.95 confidence bounds and percent sampling error. The important thing to note is the consistency between the estimates of the mean and variances when aggregated from the economic regions up to the state level and from the county level up to the economic region. One should also note the consistency of estimates from the state-wide inventory with the spatial models at the state level. Similar results were obtained for the other variables.

Discussion

This study evaluated the use of the synthetic and regression method for estimating forest stand characteristics for small geographical regions in the State of Jalisco, Mexico. The synthetic estimator can be easily applied and tends to be widely accepted due to its ease of use. However, this method will be biased if the model assumptions leading to the estimator are not satisfied, and the magnitude of this bias will likely vary with each application. The ability of the synthetic estimator to derive relatively stable estimates for small areas depends upon the extent to which the stratification scheme accounts for the local variability in the parameter of interest. In this study, this was addressed in the way the state-wide inventory was designed. The state was stratified based on the climatic variability and the spectral properties of forested and non-forested land cover types. Thus, the state shares characteristics with small areas through the spectral properties

used to stratify the state and the fact that estimates from the state-wide inventory are available for all strata.

In spite of the effort to account for the variability in forest characteristics through the stratification scheme some biases are evident. The differences are to some extent related to a precipitation gradient that runs east to west through the state (Fig. 2). Within a given climatic region changes in precipitation patterns can cause differences in the response variable. Assume that the distribution of the variable of interest is mound shaped and where the left tail represents the drier portion of the region and the right tail, the wetter portion of the region. If the small area is exclusively selected from the left tail of the distribution the estimated mean will appear to approach the overall mean resulting in an overestimation. But if the small area consists of high values, their estimated means will appear to decrease relative to the population. In the semi-arid region where the precipitation gradient is more pronounced, small areas within this region can be very extreme relative to the population; their means are the furthest from the population and subject to larger biases. In the tropical region, the precipitation gradient is less pronounced and small areas within this region differ from the population mean by only a little bit and the bias is inconsequential. This is evident when one compares the direct estimates for the three climatic regions with estimates for the small geographical regions that make up these regions.

Table 5. Small-area estimates of cubic volume at the regional, county, watershed and forest level in the State of Jalisco, Mexico.

Area	Name	Area (ha)	Volume (m^3/ha)	Variance	0.95 Lower Bound	0.95 Upper Bound	% SE
----------Economic Regions in the State----------							
1	Norte	877,349	22.8	2.93	19.3	26.2	15.0
2	Altos Norte	802,614	0.4	1.95	0.0	3.2	647.9
3	Altos Sur	661,719	2.7	2.04	0.0	5.6	104.8
4	Cienega	503,973	5.6	1.94	2.8	8.4	49.9
5	Sureste	732,339	43.4	3.11	39.9	46.9	8.1
6	Sur	598,816	48.5	3.28	44.9	52.2	7.5
7	Sierra de Amula	379,810	23.7	2.94	20.3	27.1	14.5
8	Costa Sur	696,204	62.4	3.13	58.9	66.0	5.7
9	Costa Norte	524,076	64.2	3.09	60.7	67.7	5.5
10	Sierra Ocidente	831,069	69.1	3.26	65.5	72.7	5.2
11	Valles	627,122	18.7	2.91	15.3	22.1	18.2
12	Centro	589,920	13.4	2.90	10.0	16.8	25.4
Average			31.8	2.79	28.4	35.1	10.5
Total		7,825,010					
----------Counties in Region 10----------							
1	Mascota	185,043	79.9	3.51	76.1	83.6	4.7
2	Mextlan	63,196	46.2	2.67	42.9	49.4	7.1
3	Atenguillo	59,727	49.5	2.20	46.6	52.5	6.0
4	Ayutla	88,246	48.7	3.03	45.2	52.2	7.1
5	Talpa de Allende	197,824	100.6	3.86	96.7	104.5	3.9
6	San Sebastian del Oeste	110,759	75.8	3.26	72.2	79.4	4.8
7	Cuautla	42,778	56.4	2.30	53.4	59.5	5.4
8	Guachingo	83,495	21.1	2.96	17.7	24.5	16.3
Average			69.1	3.22	66.5	72.7	5.2
Total		831,069					
----------Small Areas of Interest----------							
Watersheds 1 and 2		8,419	84.8	3.22	81.2	88.3	4.2
Forest		654	95.4	4.70	91.0	99.7	4.5
----------State-Wide Estimates----------							
State-wide Inventory		7,825,010	31.7	2.28	28.7	34.7	9.5
Spatial Model		7,825,010	31.0				

The regression method demonstrated a significant improvement over the synthetic estimator through the use of the spatial models to geographically reflect differences in the variable of interest. The accuracy of the estimates relies heavily on the quality of the spatial models; the stronger the correlation between the spatial models and the variable of interest, the less potential bias. This correlation is, to some extent related to the spatial resolution of the spatial models. Models with a fine spatial resolution are expected to be more informative than those based on a coarser resolution. If the resolution of the spatial model does not match the resolution of the field data this could introduce random noise into the estimation process and result in biased estimates because the spatial models are not able to account for the local variability in the small areas. The regression estimators were also, for the

most part, more precise than the synthetic estimator. In estimating the variances, no attempt was made to take into consideration the uncertainty associated with estimates from the spatial models. It is not uncommon to assume such estimates are error free, and therefore not account for their effects in the model and variance calculations.

Our method of small area estimation may not be compatible with certain types of data inquiries. Recently, Reich et al. (2008a) developed species distribution maps of several important tree species using the topographic and climatic variability for the State. Accuracy assessments indicated the models overestimated species occurrences compared to the state-wide inventory. The disparity between the two estimates could be attributed to the methods used to develop the maps as well as the survey itself. First, the maps did not necessarily indicate where a tree species occurred, but rather where they had a predicted likelihood of being found. The climatic data may have been emphasizing the fundamental niche (Kearney and Porter 2004) of the tree species rather than their realized niche. Second, the state-wide inventory was not designed as a vegetation survey and given the diversity of the forests and the low sampling intensity associated with certain resource characteristics, it is not surprising that the models overestimated species occurrences. If such models are used for small area estimation, one must be aware of the implications of the discrepancies between the state-wide inventory and the spatial models due to false-positives and false-negatives errors and how they may impact the estimation process.

State and local agencies need valid and relatively precise data to account for estimates of important ecological indicators that are fundamental to support decision making processes for ecosystem sustainability. Model building takes time and effort and it may not be possible to develop reliable models for every situation. However, there may be a set of spatial models which are related to the variable of interest and could potentially be used to obtain improved estimates. The decision to use this approach would depend on the statistical validity of the estimates as well as their acceptance by the policymakers who intend to use them. The fundamental issue is whether or not the small geographical regions are individually unique and thereby require estimates based solely on their own separate data. If this is the case, then only direct estimates are appropriate, and a sample survey should be designed accordingly. It would therefore be logical for the state to either coordinate outside efforts examining other variables, or alternative statistical methods to provide small area estimates for some of the more important indicator variables. This would ensure that efforts are not duplicated and are maximally cost-effective.

Conclusions

In summary, this regression-based small-area estimation procedure is relatively computationally simple and provides a significant improvement over the often used synthetic method. This study provided evidence that IMRENAT can be used as a valid source of data. The small area estimation methodology presented in this study produces consistent estimates of forest stand structure at multiple spatial scales. By applying regression models derived from the state level to more localized levels, this method allows the creation of small area estimates for many different geographical regions within the state. The resulting data are informative for decision making and planning efforts to expand assistance to land owners in managing their natural resources. We anticipate that this will improve planning among local agencies with limited resources.

References

Cochran WG. 1977. *Sampling techniques*. New York: John Wiley & Sons, 428 pp.

DuMouchel WH, Duncan GJ. 1983. Using sample survey weights in multiple regression analysis of stratified samples. *J. American Statistical Association*, **78**: 535−543.

Fay RE, Herriot RA. 1979. Estimation of income for small places: an application of James-Stein procedures to censis data. *J. American Statistical Association*, **74**: 269−277.

Flores-Garnica JG, Mendoza-Briseño MA, Aguirre-Bravo C. 2007. Monitoreo de ecosistemas con estrategias geostadisticas, una aplicación de gran escala en Jalisco, Mexico. *Madera y Bosques*, **13**: 97−104.

Ghosh M, Rao JNK. 1994. Small area estimation: an appraisal. *Statistical Sciences* **9**: 55−93.

Kearney M, Porter WP. 2004. Mapping the fundamental niche: physiology, climate and the distribution of a nocturnal lizard. *Ecology,* **85**: 3119−3131.

Laake P. 1978. An evaluation of synthetic estimates of employment. *Scandinavian Journal of Statistics*, **5**: 57−60.

Lehtonen RP, Pahkinen E. 2004. *Practical Methods for Design and Analysis of Complex Surveys*. Chichester, England: John Wiley & Sons, Ltd., 360 pp.

Pfeffermann D. 1999. Small area estimation – Big developments. Keynote Paper, Conference on Small Area Statistics, Riga, Latvia, August 1999.

Rao JNK. 1999. Some recent advances in model-based area estimation. *Survey Methodology*, **25**: 175−186.

Rao JNK. 2003. *Small Area Estimation*. Hoboken, New Jersey: John Wiley & Sons, 313 pp.

Reich RM, Aguirre-Bravo C, Bravo VA. 2008a. New approach for modeling climatic data with applications in modeling tree species distributions in the States of Jalisco and Colima, Mexico. *Journal of Arid Environments*, **72**: 1343−1357.

Reich RM, Aguirre-Bravo C, Mendoza-Briseño MA. 2008b. An Innovative approach to inventory and monitoring of natural resources in the Mexican State of Jalisco. *Journal of Environmental Monitoring and Assessment*, **146**: 383−396

SEDER-FIPRODEFO. 2007. *Inventario y Monitoreo de los recursos Naturales del Estado de Jalisco-Reporte 2006*. Gobierno del Estado de Jalisco, Secretaria de Desarrollo Rural (SEDER), Fideicomiso para la Administración del programa de Desarrollo Forestal (FIPRODEFO), Guadalajara, Jalisco, Mexico., Schreuder HT, Gregoire TG, Wood GB. 1993. *Sampling methods for multiresource forest inventory.* New York: John Willey and Sons, Inc. 464 pp.